高速なネットワーク・システムを構築するためのメモリ選択

——レイテンシ重視にはQDR-II SRAMを，容量重視にはRLDRAM-IIを

デバイスの記事　システムの記事

Anuj Chakrapani

高速ネットワーク・システムを構築する上で欠かせないメモリについて説明する．こうしたアプリケーションに特化したメモリ技術が開発されているが，ここでは特にQDR SRAMとRLDRAMについて特徴をまとめる．また，ライン・カードを例に，システム要件に応じたメモリを選択する際の視点を解説する．（編集部）

今日の高速ネットワーク・アプリケーションでは，広帯域・大容量のメモリが要求されます．例えば，標準的なネットワーク・ライン・カードでは，多くの機能のうち，パケット・バッファ，テーブル・ルックアップ，およびキュー（待ち行列）管理など，さまざまなオペレーションにメモリが必要です．メモリの帯域幅がアプリケーションのスループットに対してボトルネックにならないよう，適切にメモリを選択することが極めて重要になります．ネットワーク・アプリケーション向けに現在提供されているメモリとしては，NoBL（No Bus Latency）SRAM，QDR（Quad Data Rate）SRAM，RLDRAM（Reduced Latency dynamic RAM）などが挙げられます．

本稿では，この分野で広く使用されているQDR SRAMとRLDRAMについて解説します．

1. QDR SRAMの特徴

QDR SRAMは，ネットワーク・アーキテクチャにおいて高性能という要求を満たすために設計された同期式SRAMです．ライン速度が数十Gbpsに達すると，従来の同期式SRAMは増大する帯域幅の要求に対応できませんでした．例えば，標準同期式SRAMは，リードとライトの間にダミー（動作していない）サイクルを必要としました．その後に登場したNoBL SRAMは，リードとライトの同時動作に対応していませんでした．

従来のメモリ技術を改善し，さまざまなネットワーク・システムに適したメモリを目指して，QDR Consortium[注1]によりQDR SRAM（QDR-I，QDR-II，QDR-II＋を含む）が開発されました．QDR SRAMの動作周波数は最大400MHz，容量は9Mビット～72Mビットに対応しています（将来的には288Mビット以上に拡張できることを見込んでいる）．

QDR SRAMは，リードとライトに独立したポートを備えているため，バスが競合する心配がなく，リードとライトを同時に動作できます．また，QDR SRAMの一つのポート上では，クロックの立ち上がりと立ち下がりの両方のエッジ（DDRインターフェース）でデータのやりとりを行います．こうして，QDR SRAMは初期の同期式SRAMと比べて，全体として理論上4倍の帯域幅が得られます．

DDR SRAM（DDR-II，DDR-II+を含む）はQDR SRAMと似ていますが，独立したリード/ライト・ポートを備えていないという点が大きな違いです．つまり，DDR SRAMはリードとライトを同時に実行できません．

● QDR-IIとQDR-II+の違い

QDR-I SRAMはQDR SRAMの最初の製品で，最大動

注1：米国Cypress Semiconductor社，米国IDT社，NEC，ルネサス テクノロジ，韓国Samsung Electronics社で構成される．

KeyWord　NoBL SRAM，DDR SRAM，QDR-II SRAM，QDR-II+ SRAM，バースト長，レイテンシ，RLDRAM-II，ライン・カード

作周波数は167MHzでした．次に登場したQDR-II SRAMは最大278MHz，さらにSDR-II+ SRAMは最大400MHzに改良されました．ここでは，QDR SRAMの中でも特にQDR-II SRAMとQDR-II+ SRAMに焦点を当てます．

図1と**表1**に，QDR-II SRAMとQDR-II+ SRAMの相違点をまとめます．内部ブロックについては，次の2点が異なります．

- **C/$\overline{C}$（差動入力クロック）：QDR-II SRAMのみ**
- **データ有効ピン（QVLD）：QDR-II+ SRAMのみ**

表1　QDR-II　SRAMとQDR-II+ SRAM の相違点

	QDR-II	QDR-II+
動作周波数（DLLあり）	119MHz～300MHz	300MHz～400MHz
構　成	×8，×9，×18，×36	×9，×18，×36
V_{DD}	1.8V±0.1V	1.8V±0.1V
V_{DDQ}	1.8V±0.1Vまたは 1.5V±0.1V	1.5V±0.1V
レイテンシ（リード遅延）	1.5サイクル	2.0，2.5サイクル
入力クロック（K/$\overline{K}$）	シングル・エンド	シングル・エンド
出力クロック（C/$\overline{C}$）	あり	なし
A0（DDR B2）	あり	なし
A0，A1（DDR B4）	あり	なし
エコー・クロック	1ペア	1ペア
パッケージ	165ピンFBGA	165ピンFBGA
個別バイト・ライト（$\overline{BWa}$，$\overline{BWb}$）	あり	あり
QVLD	なし	あり

QDR-II SRAMはC/$\overline{C}$（差動入力クロック）を備えています．このポートは，SRAMからメモリ・コントローラへのデータのラッチに使用します．C/$\overline{C}$を使用する場合（デュアル・クロック・モード動作時注2），その出力タイミングはK/$\overline{K}$（差動入力マスタ・クロック）よりも優先されます（QDR-II+ SRAMでC/$\overline{C}$がサポートされない理由は後述）．

QDR-II+ SRAMは，データ有効ピン（QVLD）を備えています．これは，メモリ・コントローラが有効データの転送された時点を認識するために使用する出力ピンです．QVLDは有効データがリード・ポートに現れる前の半サイクルの間“H”になり，最後のデータが終了する前の半サイクルの間“L”になります．400MHzという高い周波数でSRAMからのデータをラッチすることは，非常に難しいと考えられます．そのような高速ではタイミングのマージンが非常に小さいからです．このようなとき，QVLDを用いてデータ有効のタイミングを認識するわけです．

QDR-II+ SRAMのレイテンシ（読み出しコマンドが出されてから実際にメモリのデータが出力されるまでの遅延）は，2.0サイクルまたは2.5サイクルです（**図2**）．

QDR-II SRAMとQDR-II+ SRAMのI/O規格は，1.5Vまたは1.8VのHSTL（high speed transistor logic）です．

注2：これに対してK/$\overline{K}$のみを使用している場合，シングル・クロック領域で動作していることになる．

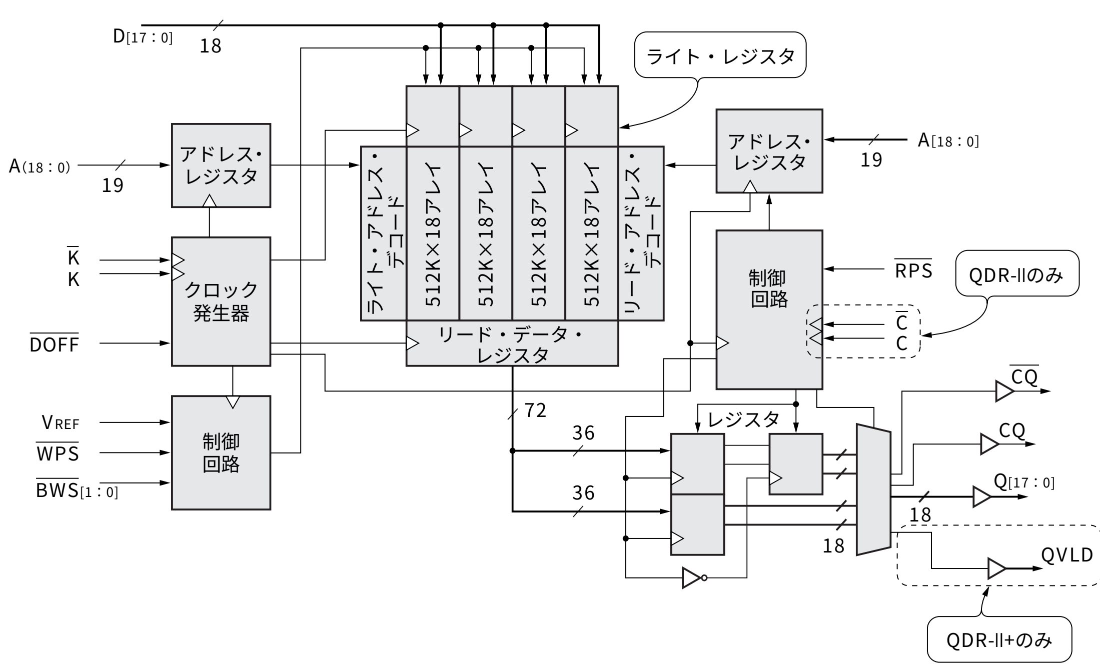


図1
QDR-II/QDR-II+ SRAM のブロック図
米国Cypress Semiconductor社のQDR-II/QDR-II+SRAM（36Mビット，バースト長4）のブロック図を示す．制御信号$\overline{WPS}$（ライト・ポート選択），$\overline{RPS}$（リード・ポート選択）および$\overline{BWSx}$（バイト・ライト選択）は，K/$\overline{K}$（差動入力マスタ・クロック）によって制御される．

リードとライトはバースト転送で行われます．バースト長は2または4に対応しています(**図3**，**図4**)．

両SRAMともに，出力ドライブ能力の設定を変更できます．また，ドライバとシステム基板のインピーダンス・マッチングが取れるように，メモリの出力インピーダンスを35～70Ωの範囲で調整することもできます．

● タイミング・マージンを考慮したクロッキング方式

メモリ・コントローラがSRAMからデータをラッチするには，次のような方法があります．

- マスタ(K/$\overline{K}$)クロックを使用
- C/$\overline{C}$を使用(QDR-IIの場合のみ)
- SRAMからのエコー・クロックを使用

低い動作周波数では，メモリ・コントローラは出力にK/$\overline{K}$クロックを使用することができます．しかし，250 MHzを超える動作周波数では，タイミング・マージンが非常に小さくなり，入力データをK/$\overline{K}$でラッチすることはとても難しくなります．そのため，QDR-II＋ SRAMではC/$\overline{C}$をサポートしていません．

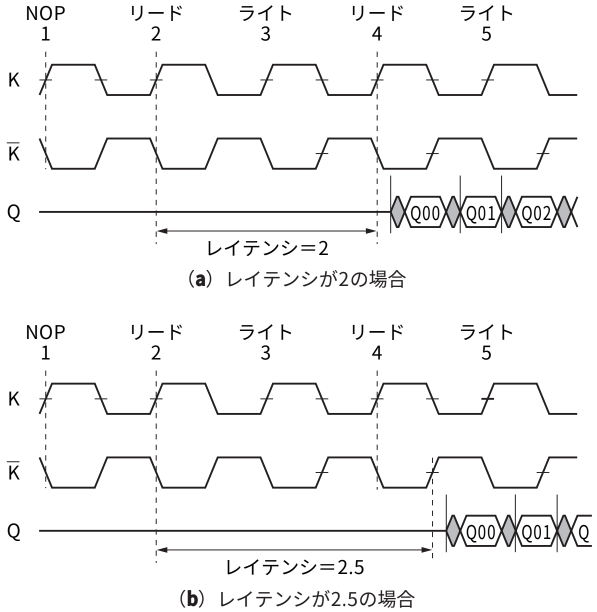


図2　QDR-II+ SRAMのリード・レイテンシ

(a)はレイテンシ＝2.0の，(b)はレイテンシ＝2.5のQDR-II+ SRAMのタイミング図．例えば，(a)ではリード開始(Kの立ち上がりクロック・エッジ2)の2サイクル後のKの立ち上がりクロックでデータ(Q00)が出力される．

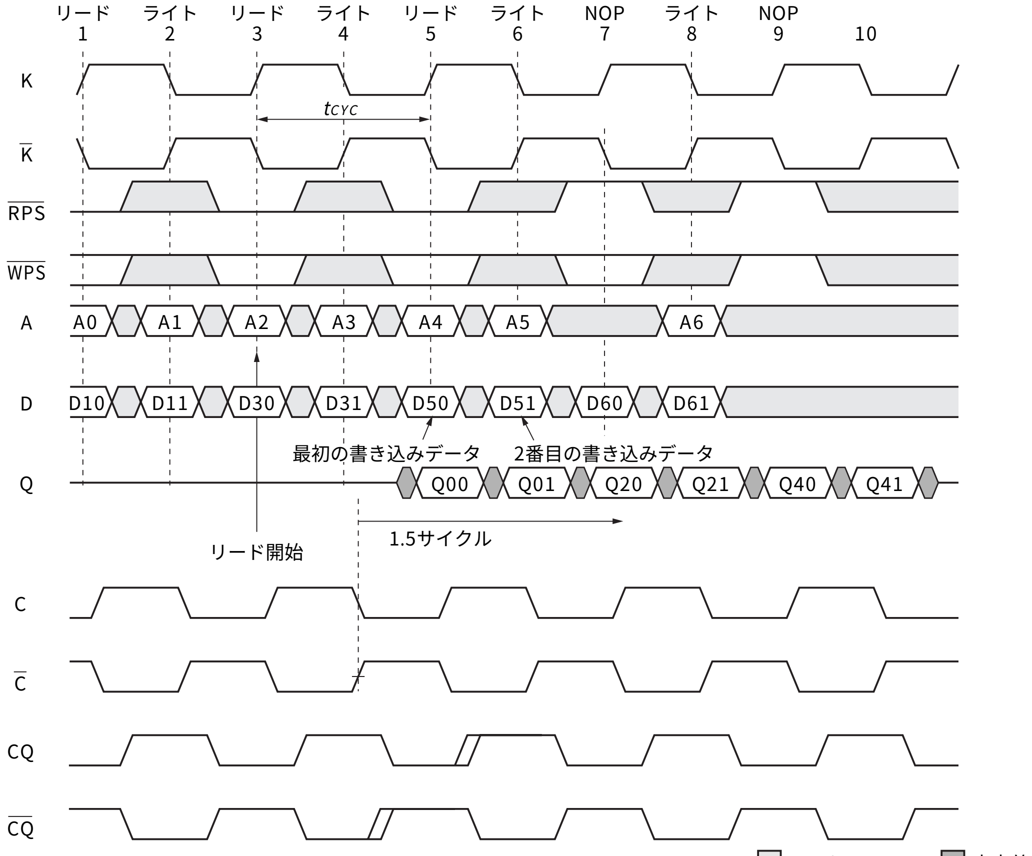
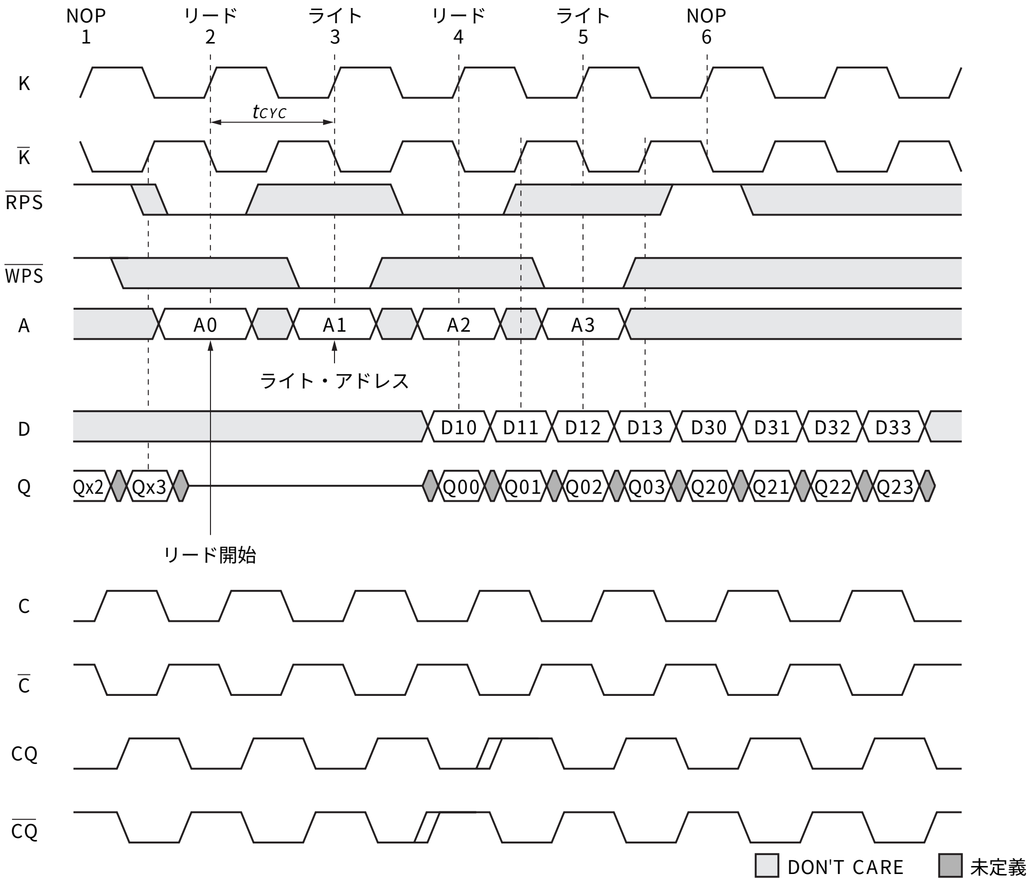


図3
QDR-II SRAMのリード/ライトのタイミング図(バースト長＝2)

$\overline{RPS}$をアサートしている間，Kの立ち上がりクロック(エッジ3)でリード・アドレスA2が与えられ，リード開始となる．そして，$\overline{C}$(シングル・クロック領域の場合$\overline{K}$)の立ち上がりから1.5サイクル後に，最初のデータ(Q20)が読み出される．書き込み時，$\overline{WPS}$は最初のデータ(D50)といっしょにKの立ち上がりエッジでアサートされる．$\overline{K}$の次の立ち上がりエッジで書き込まれる2番目のデータ(D51)といっしょにライト・アドレス(A5)が与えられる．

図4
QDR-Ⅱ SRAMのリード/ライトのタイミング図（バースト長＝4）
バースト2のデバイス（図3を参照）と異なり，リードとライトのアドレスはクロックKの立ち上がりエッジだけで与えられる．書き込み時，ライト・アドレス（A1）は$\overline{WPS}$をアサートしている間，Kの立ち上がりエッジで与えられる．最初のデータ（D10）は，次のKの立ち上がりエッジで与えられ，続く$\overline{K}$とKの立ち上がりエッジで残りの3データが与えられる．読み出し時，リード・アドレス（A0）は$\overline{RPS}$をアサートしている間にKの立ち上がりエッジで与えられる．その1.5クロック・サイクル後，最初のデータ（Q00）が$\overline{C}$（シングル・クロック領域の場合$\overline{K}$）によって読み出される．C（またはK）の次の立ち上がりエッジで2番目のデータ（Q01）が読み出される．その後，$\overline{C}$（または$\overline{K}$）とC（またはK）の立ち上がりエッジで残りの2データが読み出される．すべてのデータは，エコー・クロック（CQ/$\overline{CQ}$）とエッジがそろっている．



そのような高い動作周波数では，エコー・クロック（CQ/$\overline{CQ}$）を使用することが推奨されています．エコー・クロックは，QDR-Ⅱ SRAMやQDR-Ⅱ+ SRAMからの出力であるクロック・ペアです（シングル・エンド）．CQ/$\overline{CQ}$の周波数は入力クロックと一致し，出力データとエッジがそろっています．つまり，CQ/$\overline{CQ}$と出力データとの間には明確な位相関係が保たれるわけです（**図5**）．

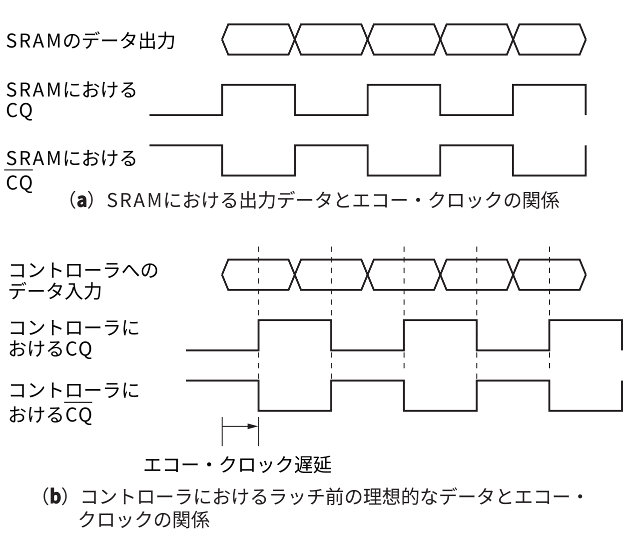


図5　SRAMとコントローラの間のエコー・クロックと出力データの関係

図6に，一つのメモリ・コントローラと二つのQDR-Ⅱ SRAMを含むシステムのブロック図の1例を示します．エコー・クロックはSRAMによって駆動されます．コントローラに出力データをラッチする前に十分なセットアップ/ホールド時間を確保するため，エコー・クロックを遅延させる（データの中心とエコー・クロックのエッジをそろえる）必要があります．正確な遅延時間は，データ有効ウィンドウ（バス上のデータが有効であることが保証された時間）に基づいて計算する必要がありますが，これは動作周波数によって変わります．次式を使用して遅延時間を計算することを推奨します．

$$t_{echo_clock_delay} = t_{cyc}/4 \quad \cdots\cdots(1)$$

エコー・クロックを遅延させるには，次のような方法があります．

- **コントローラ内のオンチップ遅延回路を利用**

メモリ・コントローラでFIFO（first-in first-out）のよう

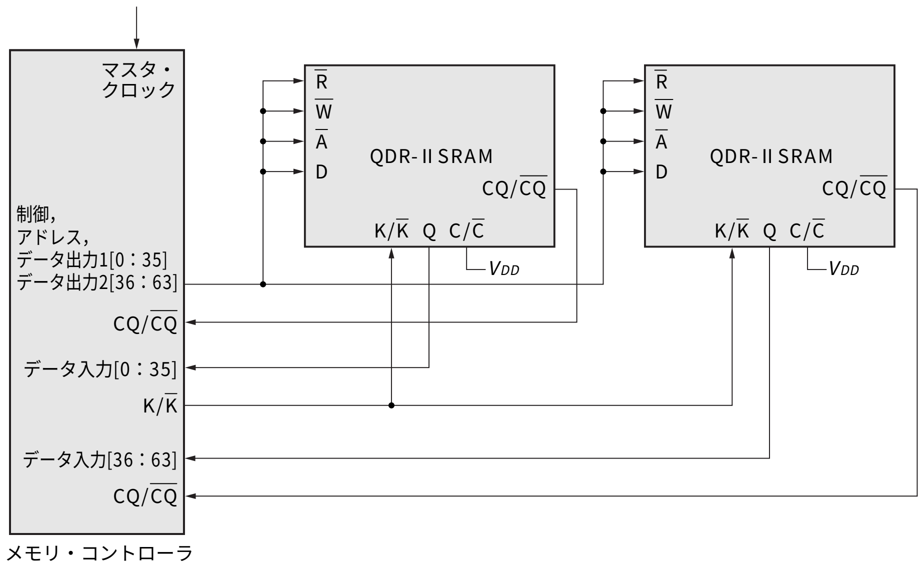


図6
QDR-II SRAMを含むシステム例
二つのQDR-II SRAMからの出力データをラッチするためにエコー・クロックを使用するコントローラを示す.

な遅延要素を使用して，エコー・クロックを遅延させます．この方法は，メモリ・コントローラがASICの場合とFPGAの場合の両方で使用できます．

●システム基板の配線遅延を利用

動作周波数に応じて，システム基板の配線遅延を利用してエコー・クロックを遅延させることができます．CQ/$\overline{\text{CQ}}$と出力データ（Q）の間の配線遅延が所望する遅延と等しくなるように，CQ/$\overline{\text{CQ}}$の配線を引き回します．特定の動作周波数の下に得られた遅延であるため，メモリ・コントローラがASICである場合に適用し，FPGAの場合に使うことはありません（通常，FPGAを搭載した基板はさまざまな周波数で動作することが要求されるため）．

●FPGA内蔵のDLLまたはPLLを利用

メモリ・コントローラをFPGAで実現する場合，FPGAが内蔵するDLL（delay-locked loop）またはPLL（phase-locked loop）を使用して，エコー・クロックを約90 °遅延させます（遅延はFPGAのセットアップ/ホールド時間によって異なる）．エコー・クロックのうちCQまたは$\overline{\text{CQ}}$のいずれか一つだけを使用して，これを90 °，270 °と遅延させて二つのバースト・データをキャプチャすることもできます．しかし，同じクロックを持つ二つのバースト間の位相関係は保証されないため，この方法は推奨できません．

これらの方法を用いる際に注意することがあります．同一のサイクルの中で，エコー・クロックとデータ出力は厳密に一致しています．再同期時，エコー・クロックが半サイクル以上遅延してしまうと，K/$\overline{\text{K}}$とエコー・クロックの間のジッタ関係が変化し，その結果，データ有効ウィンドウが狭くなって誤った結果が得られる可能性があります．

2. RLDRAMの特徴とメモリ選択

次に，同じネットワーク・アプリケーション向けのメモリであるRLDRAMについて説明します．標準的なDRAMは，ランダム・サイクル時間（t_{RC}）の影響を受けます．t_{RC}は同じ行，あるいは同じバンク[注3]への連続するアクセスを制限する時間です．RLDRAMは，ネットワーク・アプリケーション向けにt_{RC}の影響を低減するよう改善されたメモリ・アーキテクチャになっています．

RLDRAM-IIは，「ラウンドロビン・アドレス」と呼ばれる，内部バンクに順次アクセスする方式を使用して良好な帯域幅を実現しています．ただし，データ・アクセスがランダムである場合はそれほど効果的ではありません（**図7**）．データ・パターンが予測不能である場合，短いランダムなバーストの際にt_{RC}遅延が帯域幅に影響を及ぼします．**図7（a）**の例では，バンクAへの2回目のアクセスは，t_{RC}が経過するまで数サイクル待つ必要があり，その結果データ・バスは利用されない状態になります．

図7（a）に示す通り，RLDRAM-IIは最初のレイテンシが6サイクルです．短いデータ・アクセスが連続して行われる場合は，この初期レイテンシが問題になります．また，

注3：バンクは，DRAMのメモリ・アレイ（行×列）のこと．

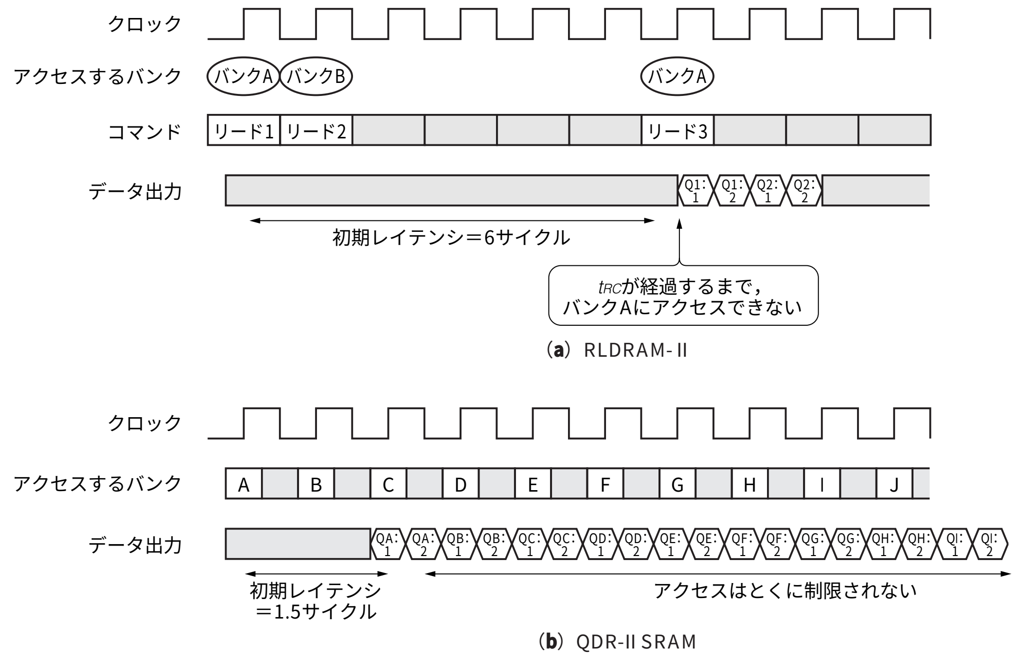


図7
RLDRAM-IIとQDR-II
(a)はバースト長が2，動作周波数が250MHzのRLDRAM-IIのデータ・アクセスを示す．読み出しコマンドが出てからのレイテンシは6サイクル．(b)はQDR-II SRAMのデータ・アクセス(バースト長と動作周波数はRLDRAM-IIと同じ)．レイテンシは1.5サイクル．

図7(a)は，短いバースト長によってRLDRAM-IIの帯域幅利用が制限される可能性も示しています．バースト長が長いと，新たなアクセスのためにより多くの時間を利用できます．従って，バースト長が長いデバイスを使用した場合，t_{RC}遅延をカバーするために交互にアクセスしなければならないバンク数は少なくて済みます．しかし，バースト長が短いデバイスでは，最大の帯域幅を得るため，より多くのバンクを交互に使用しなければなりません．

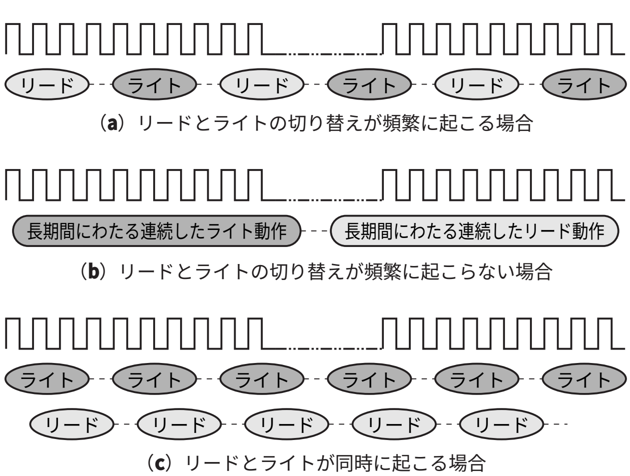


図8　リード，ライトのパターン
システムがどのようなタイミングでリードやライトを行うかによって，メモリを選ぶことが重要．

● 正しいメモリの選択が大切

アプリケーションに適したメモリを選ぶ際，バス使用率を考慮する必要があります．

システムによっては，リード時間とライト時間が短く(ショートターム)，かつそれぞれの回数が等しい(リード/ライト比＝1：1)場合，頻繁にインターリーブ(バンクに並行に交互にアクセス)することがあります(**図8(a)**)．このようなケースでは，リードからライト，ライトからリードへのバスの切り替えにダミー・サイクルを挿入することで，多くのサイクルが無駄になり，帯域幅に悪影響を与えます．こうしたシステムでは，データ・ポートがセパレートI/OのDDR-II SRAM[注4]やRLDRAM-II，あるいはQDR-II SRAMが選択肢として有力です．

一方，データ・ポートがコモンI/OのDDR-II/DDR-II+ SRAMやRLDRAM-IIは，長いバーストの中でリード動作とライト動作が行われ(ロングターム)，リード/ライト比が1：1であるアプリケーションに適していると考えられます(**図8(b)**)．長いシーケンスの中でリードとライトが頻繁に切り替わることなく行われる場合，バスの競合を回避するために無駄になるサイクル数は，リード動作やライ

注4：DDRデバイスには，コモンI/O(データの入出力が同一ポート)とセパレートI/O(データ入出力が別々のポート)の2種類がある．どちらのデバイスも，ある一つの時間にリードとライトの一方しか実行できないが，入力と出力のポートが異なるセパレートI/Oはバスの切り替えのためにサイクルが無駄になることはない．

ト動作に使用されるサイクル数と比べて非常に少なくなります．このようなアプリケーションでセパレートI/Oのデバイスを選択すると，サイクルのかなりの部分で1組のI/Oが無駄になると考えられます．

リードとライトが同時に行われるようなシステムには，QDR-II SRAMやRLDRAM-IIのセパレートI/Oに対応したデバイスが向いています（**図8（c）**）．

● ライン・カードのためのメモリ選択

このほかにも，ネットワーク・システム設計者がメモリを選択するにあたって着目するべき点はあります．例えばライン・カードの場合，さまざまな機能（ルックアップ・テーブル，パケット・バッファ，キュー管理など）のために複数のメモリが必要です．これらの機能はどれも高性能のメモリを必要としますが，すべての高速ネットワーク用のメモリが適しているわけではありません．ここでは，ライン・カードのさまざまなメモリ要件を洗い出します．標準的なライン・カードの構成を**図9**に示します．

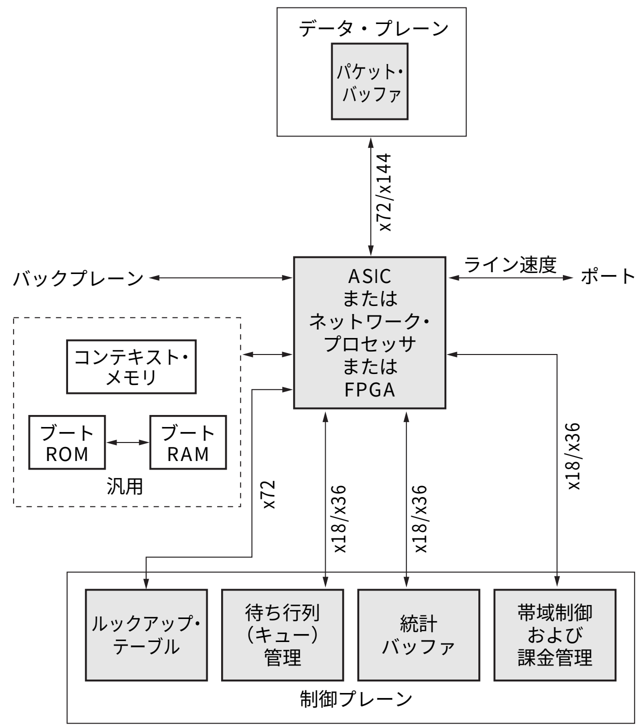


図9　標準的なライン・カードの構成要素

1）ルックアップ・テーブル

ルックアップ・テーブルは，パケットをルーティングするときにアドレスの変換を行うものです．ライン・カードの制御プレーン[注5]に搭載されています．ルックアップ・テーブルへのメモリ・アクセスはランダムに行われることが多く，短いバーストのリード動作であることが特徴です．従って，この場合のメモリの選択ではレイテンシが最も重要な要素です．最近では，多数のエントリを処理するルータについては，メモリ容量も重要な要素になってきました．

こうした要件を満たすメモリとしては，レイテンシの少ないQDR SRAMなどが挙げられます．

2）キュー管理

ライン・カードのキュー管理およびフロー制御は，ランダム・アクセスのリード/ライトが特徴です．従って，予測不能なデータ・パターンに伴うレイテンシが少ないということがメモリを選択する上で考慮すべき要素です．この場合もレイテンシが少ないメモリとしてQDR SRAMが有力候補として挙げられます．

3）統計バッファ

統計バッファでは，課金や診断結果など，さまざまな情報を扱います．パケット処理中，統計データを迅速にアクセスする必要があり，低いレイテンシであることが不可欠です．しかし，統計データは通常大きくないので，短いバーストまたはバーストなしの動作となります．QDR SRAMやNoBL SRAMが，このアプリケーションに適しているといえます．

4）パケット・バッファ

データ・プレーンの中のパケット・バッファは，パケットの処理が行われている間，出力ポートおよびスイッチ・ファブリックのパケットを一時的に保存しておくものです．パケット・バッファ用のメモリは，ASICまたはネットワーク・プロセッサの処理速度に応じて非常に高速または大容量であること，あるいはその両方を兼ね備えていることが必要になります．

高速であること（レイテンシが少ないこと）が重要なシステムではQDR SRAMが，容量が最優先のシステムではRLDRAM-IIが選択肢として考えられます．

Anuj Chakrapani
米国Cypress Semiconductor社
Memory and Imaging Division

注5：制御プレーンはネットワークを制御するための通信経路．データ・プレーンはデータ通信に用いる．

column　高速メモリ・インターフェースの設計のヒント

QDR-II SRAM（動作周波数は300MHz程度）やQDR-II+ SRAM（同400MHz）のような高い周波数で動作しているデバイスでは，タイミング・マージンは極めて厳しくなってきます．そのため，クロック発生器や制御回路において高精度のタイミングが要求されるだけでなく，高いスループットを達成する際にメモリ・インターフェースがボトルネックにならないよう，適切な基板配線および終端スキームを使用しなければならなくなります．ここでは，タイミングの問題を回避するための高速インターフェースを設計する際のヒントをいくつか紹介します．

1）配線長のマッチング

厳密な入力セットアップ/ホールド時間を実現するために，入力（アドレスA，データ書き込みD，制御信号）の配線長は，入力クロックK/$\overline{K}$の配線長と一致させる必要があります．同じく，エコー・クロックCQ/$\overline{CQ}$の配線長をSRAMからのリード・データQと一致させなくてはなりません．

2）コントローラ内のマッチング遅延

メモリ・コントローラのデータ経路内でアドレスとデータの生成とで同じDDR回路を使用することによって，コントローラ内で生成される入力（アドレス，ライト・データ）と入力クロックK/$\overline{K}$の間のタイミングの不整合を最小限にすることができます．また，同じPLLからの二つのクロック（位相を相互に90 °シフト）を使用して入力クロックとアドレス，データ，制御信号を生成できれば（先行するクロックをアドレス，データ，制御信号の生成に使用，後のクロックをK/$\overline{K}$の生成に使用），遅延を一致させクロックに対し入力の中心をそろえて，セットアップ/ホールド時間の要件を満たすことができます．

3）クロック・スキューの要件

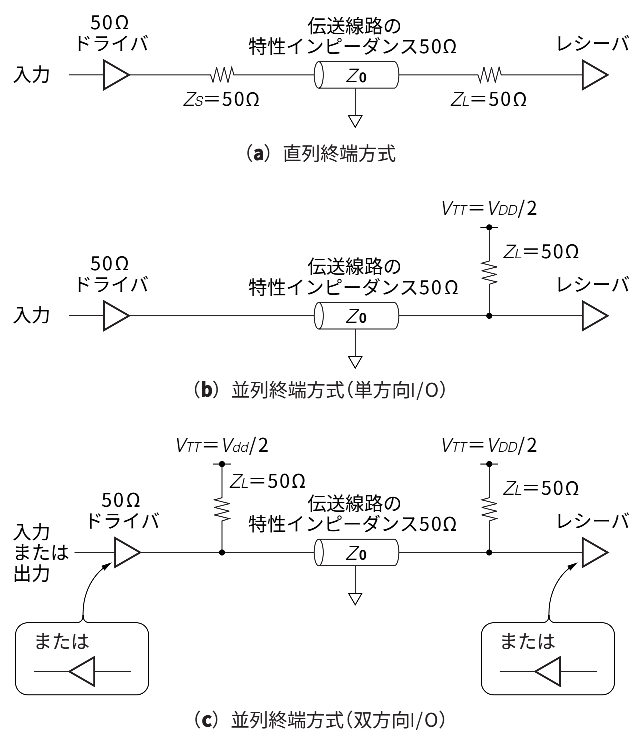
入力ジッタt_{KHKH}[注A-1]のクロック・スキューの要件を満たすため，Kと$\overline{K}$を隣接するピンに配線します．また，ノイズ耐性を高めるため，これらのピンの周囲をグラウンドで囲む必要があります．

4）終端スキーム

インピーダンスの不整合が起こると，信号が反射してリンギングが生じ，システムの信頼性を損なうことになります．リンギングが生じると，しきい値がシフトするためレシーバのダイナミック・レンジが狭くなり，不正なトリガが発生します．こうした反射をなくすため，ドライバ側の出力インピーダンスとレシーバ側の入力インピーダンス（負荷インピーダンス）を伝送線路の特性インピーダンスと一致させる必要があります．終端方式としては，次のようなものが挙げられます（**図A-1**）．

注A-1：Kの立ち上がりエッジから$\overline{K}$の立ち上がりエッジまでの時間．これは理想的にはクロック時間の1/2だが，マージンを少し持たせることができる．

- **直列終端**――ドライバ（またはレシーバ）と伝送線路の間に直列抵抗を入れ，ドライバ（またはレシーバ）のインピーダンスとこの直列抵抗の和が伝送線路のインピーダンスと等しくなるようにインピーダンスを整合させる方法．ただしこの方式の大きな欠点は，RC時定数により，立ち上がり/立ち下がり時間が劣化すること．QDR SRAMのようには出力が非常に早い速度で切り替わるものには向かない．
- **並列終端（単方向I/O）**――この方式は，レシーバ側をV_{TT}（通常V_{DD}/2に等しい）にプルアップする．プルアップ抵抗値を変えて，出力電圧幅を変えることが可能．一般に，この方式を用いると出力の“H”と“L”のしきい値はV_{TT}を中心として上下に同じレベルとなる．例えば，V_{DD}＝1.8VのQDR-IIデバイス（V_{TT}＝0.9V）では，1.35Vで“H”に，0.45Vで“L”に切り替わる．つまり出力が“L”の場合でも電力消費があり，これがこの方式の欠点である．ただし，QDR SRAMは消費電力よりも信号品質（シグナル・インテグリティ）が重要となる高速通信機器などに使用されるため，この終端方式が適している．
- **並列終端（双方向I/O）**――上記の方式に似ているが，V_{TT}へのプルアップ終端抵抗をドライバとレシーバの両方で使用する点が異なる．この方式は，同じI/Oのセットを使用して読み出しと書き込みの動作を行うDDRデバイスなどで使用される．



（a）直列終端方式

（b）並列終端方式（単方向I/O）

（c）並列終端方式（双方向I/O）

図A-1　終端方式